DOCUMENT RESUME

ED 166 249 TM 008 285

AUTHOR Lorentz, Jeffrey L.
TITLE Methodological Considerations in Teacher Effectiveness Research.
PUB DATE 27 Oct 78
NOTE 20p.; Paper presented at the conference of the Southeastern Regional Association of Teacher Educators (Biloxi, Mississippi, October 27, 1978) ; Figures 1 thru 5 are marginally legible

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; Classroom Research; *Data Analysis; *Data Processing; Discriminant Analysis; *Effective Teaching; Observation; *Research Methodology; Scoring; *Teacher Evaluation; Test Reliability

ABSTRACT

The Carroll County (Georgia) Competency-Based Teacher Certification Project developed a measure of teacher effectiveness, the 341-item Georgia Assessment of Teacher Effectiveness, based upon five previously published classroom observation systems. Three concerns were drawn from the development of this instrument: (1) the reduction of observational data to meaningful units for later analysis; (2) the scoring of these data to produce reliable indices of teacher behavior, and the provision for meaningful feedback in the form of behavioral profiles; and (3) the use of discriminant analysis to contrast a priori groups of teachers on the behavioral indices. The data were recorded by observers and keypunched to be transcribed to machine-readable form. Flexibility, ease of transcription, and compactness were useful characteristics of this process. The observers coded teacher behavior according to low-inference rating processes. The computer program then provided scores for behaviors which were already determined to be indicators of effective teaching. It provided functions for data aggregation, scoring of keys, extimation of reliability, and provision of profiles of each classroom. Discriminant analysis was used to compare the results for three groups of teachers (metropolitan, small town, and rural) and two groups of student teachers (from urban and rural universities). The two groups of student teachers were most extreme in their differences. The data are appended. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166249

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

METHODOLOGICAL CONSIDERATIONS IN TEACHER EFFECTIVENESS RESEARCH

Jeffrey L. Lorentz
West Georgia College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jeffery L Lorentz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A Paper Presented
At a Conference of the
Southeastern Regional Association of Teacher Educators

Biloxi, Mississippi
October 27, 1978

TM008285

METHODOLOGICAL CONSIDERATIONS IN TEACHER EFFECTIVENESS RESEARCH

Jeffrey L. Lorentz
West Georgia College

Introduction

This paper will treat several topics of concern to researchers who study teachers in actual school settings. These are drawn from a five-year study of teachers and students in several Georgia school districts. These topics include (a) the reduction of observational data to meaningful units for later analysis; (b) the scoring of these data to produce reliable indices of teacher behavior, and the provision for meaningful feedback in the form of profiles of behavior; and (c) the use of Discriminant Analysis to contrast a priori groups of teachers on the behavioral indices.

The paper will describe two specially developed computer programs which can be used with a variety of observation systems.

Background

The present paper is based on experiences in the Carroll County Competency-Based Teacher Certification (CBTC) project which has been looking at teacher behavior in actual classrooms for the past five years.

The CBTC project has applied five objective, systematic direct observation instruments in the course of this study. These include the Spaulding Teacher Activity Rating Schedule (STARS) (Spaulding, 1970); the Coping Analysis Schedule for Educational Settings (CASES) (Spaulding, 1969); the Observation Schedule and Record, Form 5 Verbal (OScAR 5V)(Medley, 1973); the Florida Climate and Control System (FLACCS) (Soar, Soar & Ragosta, 1971); and the Teacher Practices Observation Record (TPOR) (Brown, 1968). These five are measures of teacher behaviors, student behaviors and teacher-student interactions. They are designed to be used by especially-trained observers in on-going classrooms on six to twelve brief (10-20 minute) visits to each classroom. These instruments have been shown to provide stable, reliable measures of classroom behavior in a number of other studies as well as in the CBTC project.

One outcome of the CBTC study is the Georgia Assessment of Teacher Effectiveness (GATE) which is a composite measure derived from the five instruments cited above. GATE contains 341 sign

items which were adapted and synthesized from the 1,344 category and sign items on the five instruments. These were found to be the most reliable and stable items of the original 1,344 and appeared to represent the most important elements of the original instruments. Coker (1978) provides a complete description of GATE.

The utilization of five observation instruments of different origins in a common study, and of synthesizing a new instrument from the five, has provided the experience upon which the discussion below is based.

The Reduction of Observational Data

Several considerations are important from a methodological view point when using classroom observation instruments. First, and most important of all, the observation record form must be designed with the observer in mind; second, the form should facilitate modifications (either addition of new items or the expansion of existing items) during its developmental period; and finally, the record form should permit ease in reduction of the data to machine readable form.

Since the first and second of these considerations were of primary importance, the record forms used in the CBTC study were either mimeographed or printed by offset duplication. A machine-scorable (ie., "op-scan") form was not employed because of the relatively small number of forms needed and the prohibitive cost of special-purpose, machine-scorable forms.

Consequently, the observations were recorded on pages designed for the ease of use by the observer. These were inexpensive enough to permit re-printing with modifications as often as necessary.

A major problem in data reduction then, was one of transcribing the data to machine-readable form. This was done by keypunching the observational data into cards. The nature of the data and the problems arising there from are discussed below in some detail.

A standard punched card (ie., "IBM card") can contain no more than 80 numbers and/or letters. The typical observation instrument may have provision for 300 or more items of behavior to be recorded, along with descriptive information such as location of school, grade, class, teacher, date, observer, etc. Given these considerations it is apparent that a system of punching should be used which would be characterized by flexibility, ease of transcription and compactness. These characteristics fit the system of data transcription adopted.

Flexibility. The system ultimately designed can be used with a variety of observation instruments. Also, it permits modification of the same instruments without the need for modification of the transcription (ie., keypunching) process.

Ease of Transcription. The system has the virtue of permitting the transcriber (ie., keypunch operator) to punch records directly from the record form without the need for clerical intervention, transcription to code sheets, etc.

Compactness. Since most of the options available for coding on the observation record form will not occur during any given observation period, much of the data to be transcribed consists of blanks. If these are skipped, the remainder of the "data" to be punched is relatively small. Thus, fewer cards and fewer keystrokes are required.

The system employed for keypunching uses two types of records, a header card and a detail card. Each observation form is assigned a unique six-digit identification number. That number is used to collate the header and detail cards. Codes for district, school, teacher, observer, date and any other key information are punched in the header card. These data are not needed in the detail cards since they can be accessed by the common six-digit identification number.

The detail cards contain a series of four-digit codes which represent each item checked by the observer during the observation. Each item is provided with a unique number and these are punched in any order. As many detail cards are used as needed to contain the items coded.

Finally, a computer program is employed to "translate" these coded data into a binary positional tally. For example, cell 1105 represents the junction of teacher behavior 11 (Directing) with student behavior 05 (Question for verification). This is punched as a 1105. The computer then translates this code. Cell 1105 is the 14th "item" on GATE, and the program assigns a "1" to position ("column") 94 of the translated record. The final translated record can be any number of positions ("columns") long, and will contain all the identification data from the header card and ones for every cell or item checked and zeroes for items not recorded. Figure 1 illustrates this process which begins with

INSERT FIGURE 1 HERE

an observation form and ends with a computer-tape record 421 positions long (80 identification columns plus 341 items of observational data). These translated records can then be used in standard statistical programs, or in special purpose programs

such as the one described below.

This translation program is general-purpose. It can be used with any code structure and with any number of observational categories or signs.

Scoring the Observational Data

All of the observational systems discussed above are low-inference, that is, the observer is not required to make inference about the significance, value or relationship to other behaviors, of the behaviors being observed. It is the task of others (researchers, school boards, administrators, supervisors, etc.) to identify, a priori, those behaviors considered essential for good teachers.* The observation data can be scored to reflect these a priori behaviors.

For example, a desired behavior for student teachers may be that "the student teacher demonstrates enthusiasm for teaching and the topics being taught" (Coker, 1978). GATE contains four items which can be combined to reflect this behavior.

A computer scoring program has been devised which will provide comparable scores on any set of a priori behavioral composites or "keys". This program includes the following functions:

Data aggregation. Since the classroom (ie., the teacher) is the unit of analysis, all observation records over the sample period are sorted together by classroom. The program computes an average frequency for each item by classroom. This permits comparing classes with different numbers of observations. Since the frequency of occurrence is greater for some items than others, every item's contribution to a key may not be of equal weight. Thus each item is standardized to a mean of 0 and standard deviation of 1 (ie., z-score) so that each item contributes an equal weight to the key.

Scoring of keys. Next the program computes a series of composite key scores for each class by combining the appropriate items for each key. An item may contribute positively or negatively to a key. These key scores are then standardized across all classes so that the mean and standard deviation are 50 and 10, respectively, for all scores. Thus the group can be compared

---

*NOTE: We distinguish between those behaviors identified a priori as essential for good teaching and those which have been validated against measures of student growth. The former are merely guesses and often reflect current myths; the latter are the true competencies. See, for example Medley's (1977) review.

across classes on one key or one class can be compared across all keys.

Estimating reliability. The program provides an estimate of the internal consistency of each key using the Hoyt (1941) method outlined below:

ANOVA

Source	df	ms
Classes	C-1	a
Items	I-1	b
Residual	(C-1)(I-1)	c
Total	CI-1	

$$\alpha = \frac{a - c}{a}$$

This internal consistency estimate is important when new keys are developed.

Profiles. Finally, the program plots profiles of classes for each key or keys for each class. These are illustrated in Figures 2 and 3. The former, a profile of all classes for each key, would

INSERT FIGURES 2 AND 3 HERE

be of use to a supervisor or administrator working with teachers, as in an inservice program. The latter, a profile of one teacher's scores on all keys, could be used to provide feedback to individuals about their performance.

The option of punched scores as output is also available. These scores could be used in correlational studies with student growth measures for the same group of classrooms. An additional use of these scores is described below.

Again, this scoring program is general-purpose. It will score any observational system on any set of keys. While it is presently set to handle up to 200 classrooms, only the size of the computer restricts its ability to handle more classrooms.

Contrasting Intact Groups of Teachers

GATE observations were made during school year 1977-78 on the following different groups of student teachers and regular classroom teachers:

Group Number	Type of System	Number of Teachers	Number of GATE Observations Per Teacher
1	Metropolitan Teachers	53	12
2	Urban U. Student Teachers	33	12
3	Small Town Teachers	22	12*
4	Rural Teachers	26	16**
5	Rural U. Student Teachers	23	16**

* all made by same observer
** several teachers had fewer observations

Each teacher's data was scored to provide 16 scores on a set of behaviors identified by the teacher education faculty in an urban university as important for their student teachers. These are listed, along with some descriptive statistics in Table 1.

INSERT TABLE 1 HERE

Discriminant analysis (Veldman, 1967) was the statistical technique chosen to contrast these five groups. This multivariate statistical procedure can be viewed as an extension of the univariate F-test. Where the F-test is used to contrast two or more groups on one dependent variable, discriminant analysis is used to contrast two or more groups on more than one dependent variable.

In this instance, the groups were contrasted on the 16 behaviors to determine if the teachers in one group exhibited a set of teaching behaviors which was significantly different from the set of behaviors which was exhibited by the other groups. The means of the sixteen behaviors are plotted for each group in Figure 4. The results of the Discriminant Analysis revealed that

INSERT FIGURE 4 HERE

these profiles were in fact significantly different.

These group differences in profiles can be described by one or more discriminant functions. The groups are contrasted on the first two of these functions in Figure 5. This shows clearly that,

INSERT FIGURE 5 HERE

in terms of the first dimension, the two groups of student teachers are most extreme in their differences. The Rural and Urban teachers are most extreme on the second dimension. The disparate nature of three groups of teachers and the similarity of one group of student teachers to the small city teachers is seen in the two-dimensional plot.

Other results of the Discriminant Analysis reveal the behaviors which contribute most to these group differences. Since these data are exploratory in nature, they will not be discussed here.

Conclusion

These techniques of data reduction, scoring and group discrimination have been found to be useful tools in the study of teaching. They can be applied in a variety of settings including experimental/correlational studies, staff development programs and teacher assessment. The techniques are applicable to a variety of instruments in a variety of settings.

Thus, it is possible to apply appropriate measures of teacher effectiveness with the knowledge that the meaningful interpretation of the results is possible.

REFERENCES

Brown, B. B. The experimental mind in education. New York: Harper & Row, 1968.

Coker, J. G. A study of student teaching using direct observation. Unpublished doctoral dissertation, Georgia State University, 1978.

Hoyt, C. Test reliability estimated by analysis of variance. Psychometrika, 1941, 6(3), 153-160.

Medley, D. M. Observation schedule and record, form 5 verbal (OScAR 5V). Mimeograph, University of Virginia, 1973.

Medley, D. M. Teacher competence and teacher effectiveness. Washington, D.C.: American Association of Colleges for Teacher Education, 1977.

Soar, R. S., Soar, R. M., & Ragosta, M. Florida climate and control system (FLACCS). Gainesville, Florida: Institute for Development of Human Resources, University of Florida, 1971.

Spaulding, R. L. Classroom behavior analysis and treatment. Education Improvement Program. Duke University, Durham, N.C., 1969.

Spaulding, R. L. Spaulding teacher activity rating schedule (STARS). San Jose State University, San Jose, Ca., 1970.

Veldman, D. J. Fortran programming for the behavioral sciences. New York: Holt, Rinehart, & Winston, 1967.

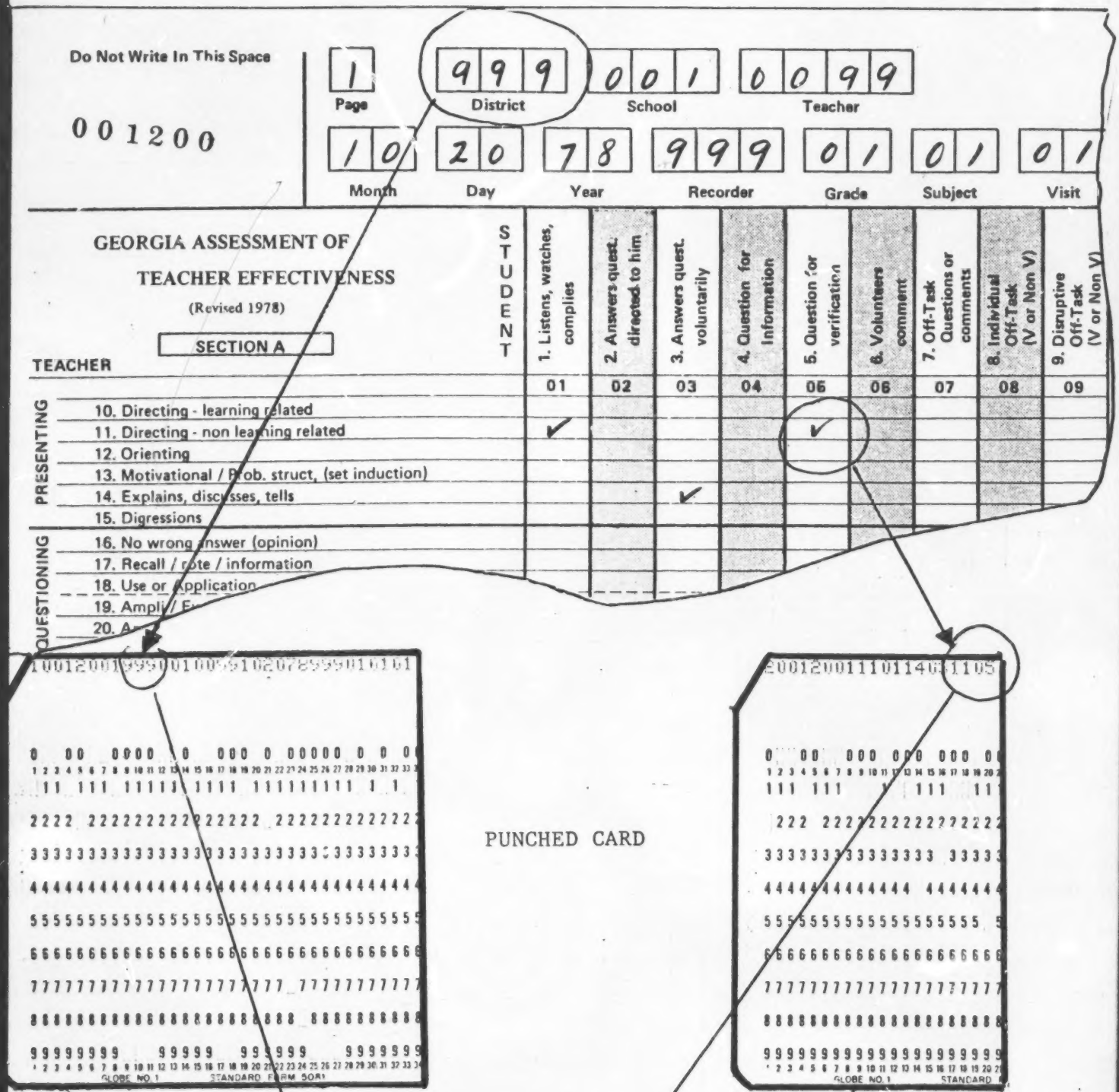
Do Not Write In This Space
001200
1 Page
999 District
001 School
0099 Teacher
10 Month
20 Day
78 Year
999 Recorder
01 Grade
01 Subject
01 Visit
GEORGIA ASSESSMENT OF
TEACHER EFFECTIVENESS
(Revised 1978)
SECTION A
TEACHER
STUDENT
1. Listens, watches, complies
2. Answers quest. directed to him
3. Answers quest. voluntarily
4. Question for Information
5. Question for verification
6. Volunteers comment
7. Off-Task Questions or comments
8. Individual Off-Task (V or Non V)
9. Disruptive Off-Task (V or Non V)
01 02 03 04 05 06 07 08 09
PRESENTING
10. Directing - learning related
11. Directing - non learning related
12. Orienting
13. Motivational / Prob. struct, (set induction)
14. Explains, discusses, tells
15. Digressions
QUESTIONING
16. No wrong answer (opinion)
17. Recall / rote / information
18. Use or Application
19. Ampli / E
20. A
PUNCHED CARD
COMPUTER RECORD
TO THIS
001200 999 001
0100101001 10 010000

FIGURE 1

TEACHER EFFECTIVENESS PROFILE

03/31/70

	SCORE
1 A-1 DEMONSTRATES ENTHUSIASM FOR TEACHING AND THE TOPICS BEING TAUGHT	
TEACHER NUMBER 76	36.66
TEACHER NUMBER 77	21.79
TEACHER NUMBER 78	31.35
TEACHER NUMBER 79	48.40
TEACHER NUMBER 80	64.86
TEACHER NUMBER 81	28.06
TEACHER NUMBER 82	54.29
TEACHER NUMBER 83	52.84
TEACHER NUMBER 84	59.02
TEACHER NUMBER 85	55.56
TEACHER NUMBER 87	38.30
TEACHER NUMBER 87	47.72
TEACHER NUMBER 89	58.58
TEACHER NUMBER 90	44.48
TEACHER NUMBER 91	65.48
TEACHER NUMBER 93	54.77
TEACHER NUMBER 94	58.86
TEACHER NUMBER 95	54.43
TEACHER NUMBER 96	47.03
TEACHER NUMBER 97	47.59
TEACHER NUMBER 98	52.61
TEACHER NUMBER 99	59.56
TEACHER NUMBER 100	47.41
TEACHER NUMBER 101	45.38
TEACHER NUMBER 102	54.43
TEACHER NUMBER 103	47.69
TEACHER NUMBER 104	57.50
TEACHER NUMBER 105	59.44
TEACHER NUMBER 106	48.51
TEACHER NUMBER 107	49.45
TEACHER NUMBER 108	56.62
TEACHER NUMBER 109	45.42
TEACHER NUMBER 120	55.20

12

FIGURE 2

13

TEACHER EFFECTIVENESS PROFILE 03/31/78

TEACHER NUMBER 10

		SCORE
1	A-1 DEMONSTRATES ENTHUSIASM FOR TEACHING AND THE TOPICS BEING TAUGHT	64.86
2	A-3 PROVIDES OPPORTUNITIES FOR SUCCESS EXPERIENCES BY STUDENTS	58.20
3	A-4 DEMONSTRATES PATIENCE EMPATHY AND UNDERSTANDING	53.78
4	B-2 IDENTIFIES LEARNING STYLES RATES OF LEARNING AND CAPABILITIES OF STUDENTS	53.86
5	B-3 DEM UNDERSTAND OF PROCESSES INVOLVED IN SELECT OF LEARN CONTENT AND METHOD	72.00
6	C-1 DETERMINES OBJ (SR) AND GOALS (LR) APPROP TO EACH STU/GROUPS IDENT NEEDS	46.66
7	C-3 SPECIFIES TEACHING PROCEDURES	66.10
8	C-4 IDENTIFIES ASSESSMENT PROCESS	59.51
9	D-1 MAINTAINS STUDENT INVOLVEMENT IN LEARNING TASKS	60.11
10	D-2 USES ACTIV'S WHICH CALL FOR PUPIL PLAN,OBSERVING,DESCRIB,EXPERIM & WRIT'G	60.85
11	D-3 ORGANIZ AND USES RANGE OF APPROP INSTRUCTIONAL MATERIALS AND EQUIPMENT	54.85
12	D-4 USES A VARIETY OF COG LEVELS IN QUESTIONING STRATEGIES	75.72
13	E-1 GIVES DIRECTIONS CLEARLY	40.91
14	E-2 MANAGES DISRUPTIVE BEHAVIOR CONSTRUCTIVELY	52.80
15	L-5 HELPS STUDENTS RECOGNIZE PROGRESS AND ACHIEVEMENTS	55.77
16	I-3 USES A VARIETY OF EVAL METHODS WHICH PROVIDE STU WITH SPEC FEEDBACK	48.98

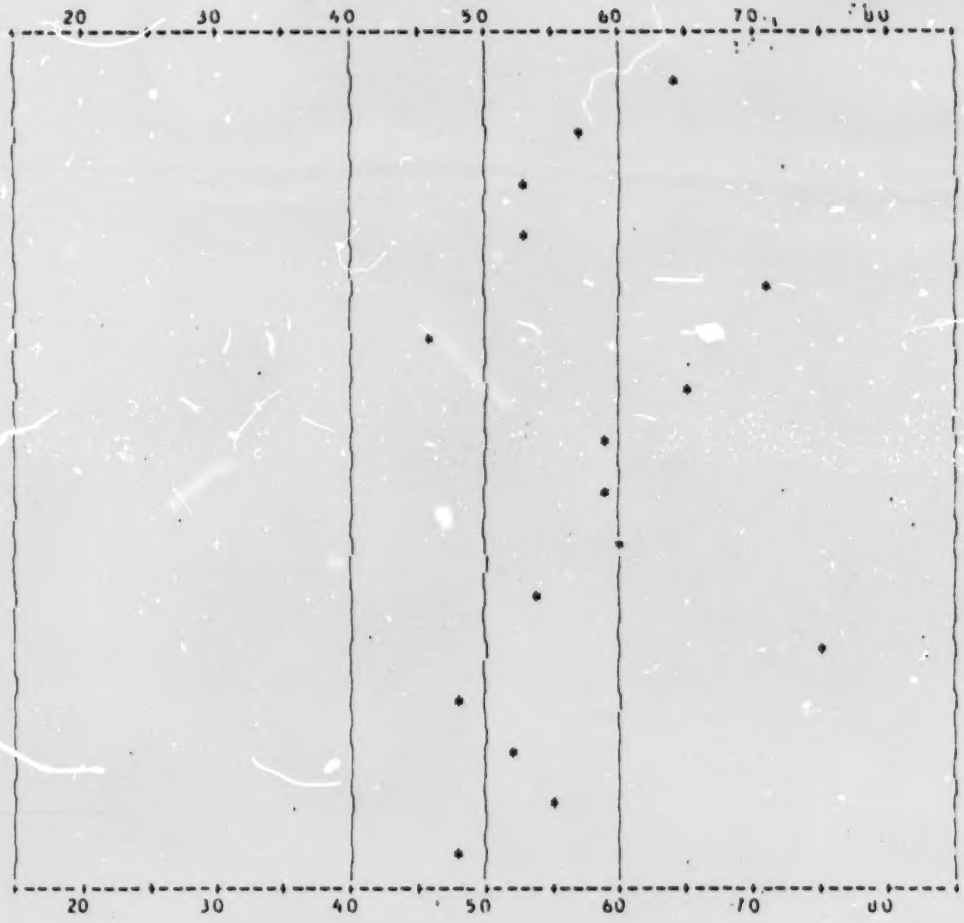

FIGURE 3

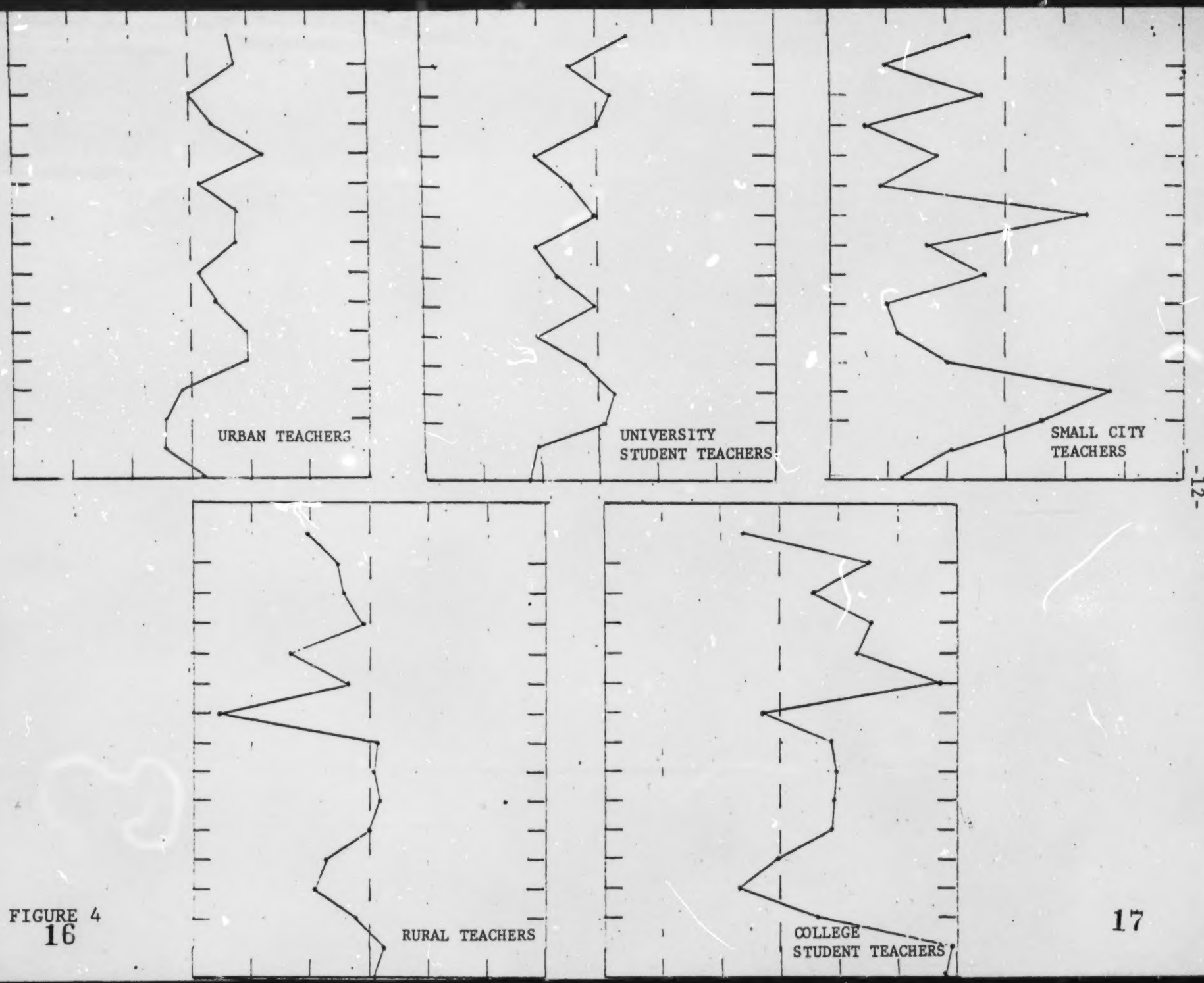
URBAN TEACHERS
UNIVERSITY STUDENT TEACHERS
SMALL CITY TEACHERS
RURAL TEACHERS
COLLEGE STUDENT TEACHERS

FIGURE 4

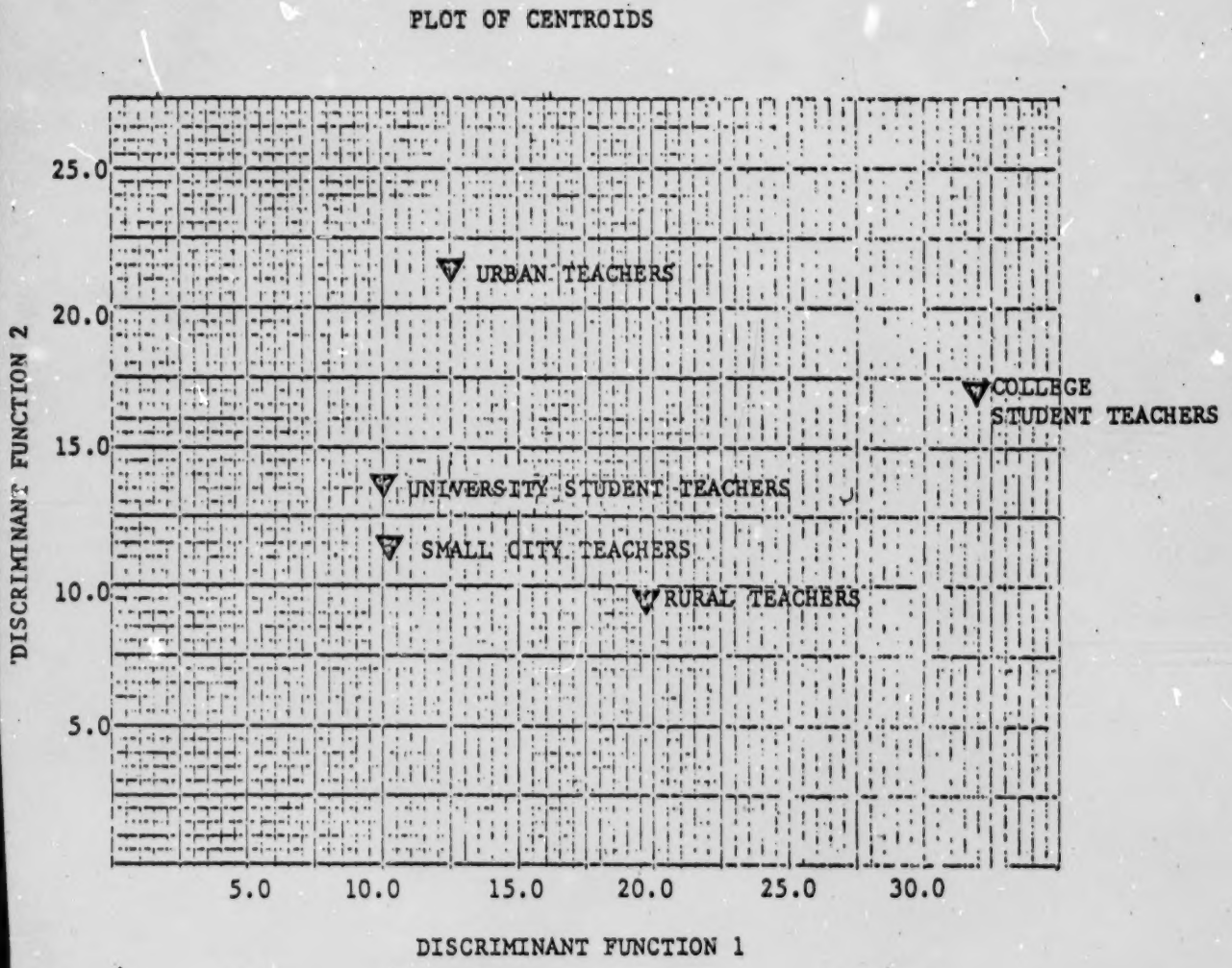
PLOT OF CENTROIDS
25.0
20.0
15.0
10.0
5.0
DISCRIMINANT FUNCTION 2
URBAN TEACHERS
COLLEGE
STUDENT TEACHERS
UNIVERSITY STUDENT TEACHERS
SMALL CITY TEACHERS
RURAL TEACHERS
5.0
10.0
15.0
20.0
25.0
30.0
DISCRIMINANT FUNCTION 1

FIGURE 5

TABLE 1

Student Teacher Behavior Keys, Reliabilities, Group Means and F-Tests

Behavior Statements	No.of GATE Items	Alpha Reliabilities	MEANS OF GROUPS 1	2	3	4	5	UNIVARIATE F-test
1. Demonstrates Enthusiasm for Teaching & the Topics being Taught	4	44	53.3	52.6	47.0	44.9	46.9	5.16**
2. Provides Opportunities for Success Experiences by Students	16	56	53.7	47.6	39.8	47.3	57.5	16.26**
3. Demonstrates Patience, Empathy, and Understanding	39	59	50.0	51.1	48.0	47.7	52.7	1.07
4. Identifies Learning Styles, Rates of Learning & Capabilities of Students	15	58	51.7	50.0	38.1	49.5	57.6	15.64**
5. Demonstrates Understanding of Processes Involved in Selection of Learning Content and Methods	18	61	56.1	44.7	44.2	43.2	56.5	22.1 **
6. Determines Objectives and Goals Appropriate to each Students'/Groups' Identified Needs	10	00	50.8	47.7	39.4	48.1	63.6	30.56**
7. Specifies Teaching Procedures	15	33	53.9	49.9	56.9	37.4	48.5	23.57**
8. Identifies Assessment Process	10	44	53.8	44.7	43.3	50.6	54.3	9.30**
9. Maintains Student Involvement in Learning Tasks	100	73	50.7	46.5	48.2	50.3	54.7	2.60*
10. Uses Activities Which call for Pupil Plan, Observing, Describing, Experimenting & Writing	20	73	52.0	49.6	40.0	50.8	54.5	8.62**
11. Organizes & Uses Range of Appropriate Instructional Material & Equipment	4	69	54.6	44.9	41.8	50.0	54.4	12.37**
12. Uses a Variety of Cognitive Levels in Questioning Strategies	14	57	54.6	48.9	45.0	46.4	49.7	5.65**
13. Gives Directions Clearly	11	26	49.2	51.3	58.9	45.4	46.5	7.74**
14. Manages Disruptive Behavior Constructively	20	48	47.9	50.5	53.0	48.9	53.1	2.02
15. Helps Students Recognize Progress and Achievements	14	78	47.9	44.9	45.4	51.2	64.9	28.16**
16. Uses a Variety of Evaluative Methods Which Provide Student with Specific Feedback	7	48	51.0	44.0	41.1	50.4	64.1	35.06**

* $p < .05$
** $p < .01$